UNIVERSITE NANCY I
IUT HENRI POINCARE DE LONGWY,
DEPARTEMENT GEA

EXERCICES CORRIGES DE STATISTIQUE DESCRIPTIVE

- TABLEAUX ET GRAPHIQUES,
- PARAMETRES (DE POSITION, DE DISPERSION, DE CONCENTRATION),
- AJUSTEMENTS (LINEAIRES ET NON LINEAIRES)

UNIVERSITE NANCY I
IUT HENRI POINCARE DE LONGWY,
DEPARTEMENT GEA

Exercices sur les tableaux, graphiques et paramètres.

EXERCICE 1 : Des enfants sont classés d'après la durée écoulée entre la date de mariage de leurs parents et la date de leur naissance. Les observations faites sont consignées dans le tableau de l'annexe. Ce tableau doit être compris de la façon suivante : dans la population étudiée, 91 enfants sont nés dans la première année de mariage de leurs parents, 72 enfants sont nés pendant la deuxième année de mariage, etc.

1) Compléter le tableau de l'annexe.
2) Interpréter la ligne « année 8 ».
3) Calculer la moyenne, la variance et l'écart-type de la variable étudiée.

EXERCICE 2 : la direction générale de l'agriculture et de la forêt nous donne la répartition par tranches d'âges des chefs d'exploitation agricole d'une région.

Moins de 25 ans	580 exploitations
De 25 à 29 ans	2 162 exploitations
De 30 à 39 ans	8 063 exploitations
De 40 à 49 ans	9 569 exploitations
De 50 à 59 ans	10 660 exploitations
Au moins 60 ans	15 913 exploitations

1) Définir la population étudiée, l'individu et le caractère ainsi que les modalités de celui-ci.
2) Compléter le tableau statistique de cette série (fréquences, fréquences cumulées croissantes et décroissantes) ; on retiendra 20 ans et 70 ans comme âge minimal et maximal.
3) Quelle est la proportion des chefs d'exploitations qui ont : au moins 40 ans ? moins de 30 ans ? entre 25 et 60 ans ?
4) Le graphique des fréquences cumulées croissantes et décroissantes est présenté en annexe Déterminer par le calcul la médiane *Mé* et les quartiles Q_1 *et* Q_3.
   Donner une estimation graphique des déciles D_1 *et* D_9. Placer les points sur le graphique.
5) Quelle est la proportion des chefs d'exploitations qui ont entre 35 et 65 ans (détermination graphique)?

EXERCICE 3: On s'intéresse à la répartition des salaires parmi les 1382 employés de l'entreprise ALPHA.

1) Compléter le tableau des données qui se trouve en annexe (les salaires sont exprimés en centaines d'euros).
2) Calculer le salaire moyen et le salaire médian. Estimer graphiquement les quartiles Q_1 *et* Q_3 et les déciles D_1 *et* D_9 (le graphique des fréquences cumulées se trouve en annexe)

   Calculer l'écart interquartile $Q_3 - Q_1$ et l'écart interdécile $D_9 - D_1$. Rappeler leur interprétation.
3) Calculer la médiale.
4) Calculer la proportion des employés dont le salaire est inférieur à la médiale Retrouver ce résultat par lecture de la courbe de concentration, donnée en annexe.
5) Calculer la part de la masse salariale totale que se partagent les employés dont le salaire est inférieur à la médiane.

EXERCICE 1

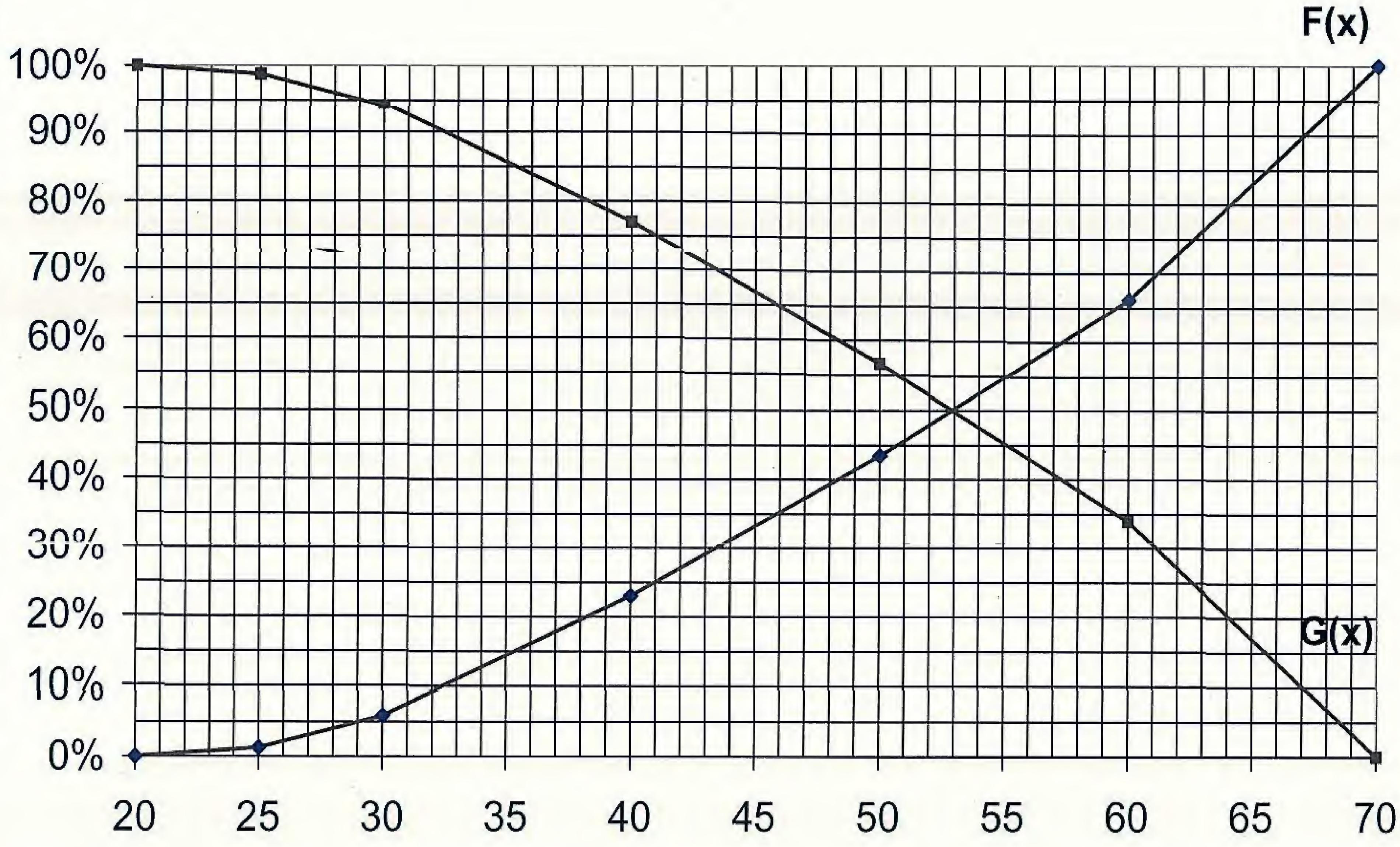
ANNEXE

Année	Nombre d'enfants	fréquence	effectif cumulé croissant	fréquence cumulée croissante	effectif cumulé décroissant	fréquence cumulée décroissante
1	91	18,2%	91	18,2%	500	100,0%
2	72	14,4%	163	32,6%	409	81,8%
3	60	12,0%	223	44,6%	337	67,4%
4	52	10,4%	275	55,0%	277	55,4%
5	45	9,0%	320	64,0%	225	45,0%
6	40	8,0%	360	72,0%	180	36,0%
7	37	7,4%	397	79,4%	140	28,0%
8	32					
9	26					
10	25					
11	20					
TOTAL:	500					

EXERCICE 2

Classes	Effectifs	Fréquences	Fréquences cumulées croissantes	Fréquences cumulées décroissantes
[20, 25[	580			
[25, 30[	2 162			
[30, 40[	8 063			
[40, 50[	9 569			
[50, 60[	10 660			
[60, 70[	15 913			
TOTAL :				

Exercice 2

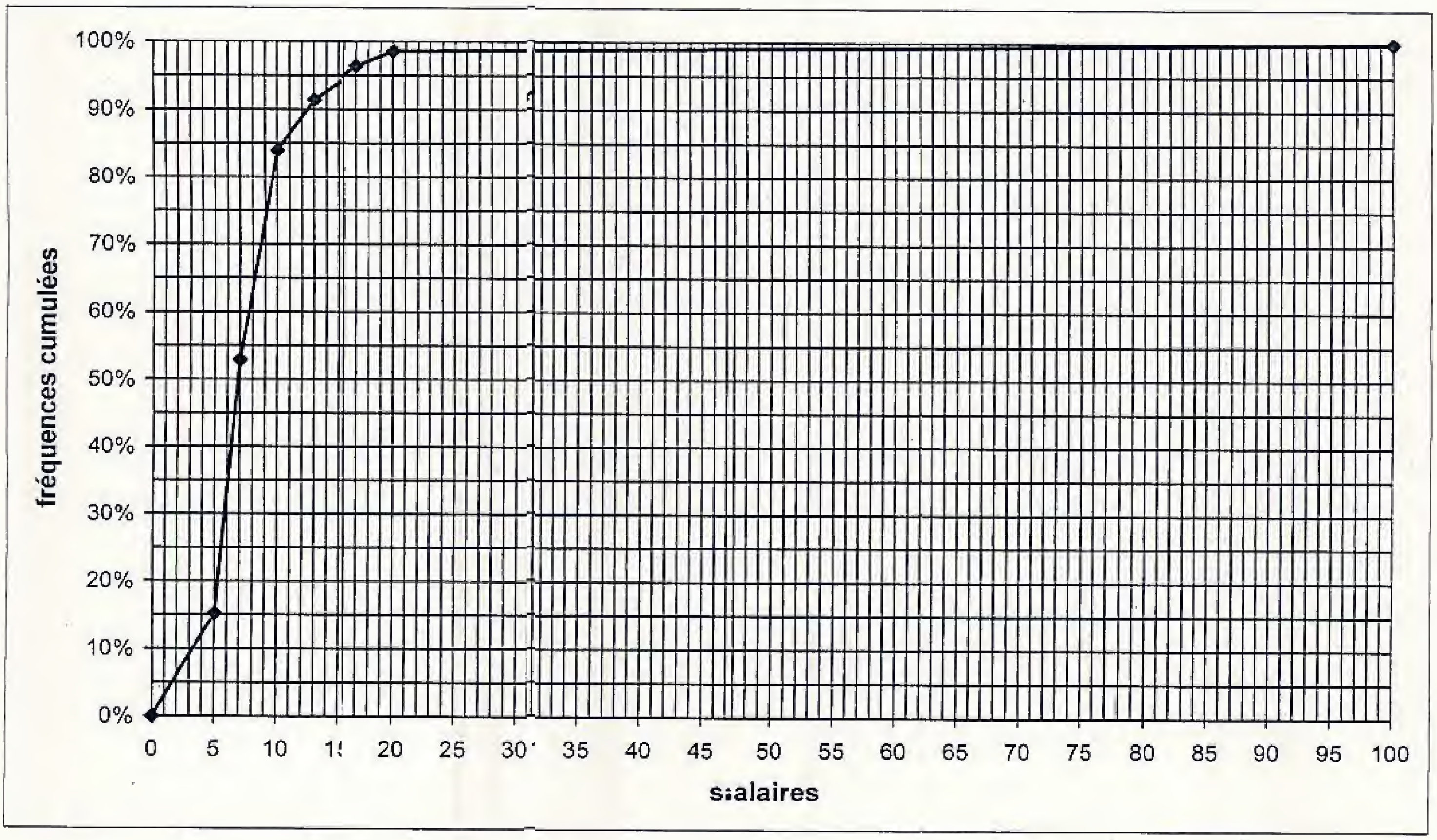


Exercice 3

Classe de salaire	centre	*effectif*	*fréquence*	fréquence cumulée	masse salariale de la classe	masse salariale cumulée	masse salariale cumulée (en %)
[0 , 5[	2,5	**210**	15,2%	15,2%	525	525	
[5, 7[	6	**520**	37,6%	52,8%	3 120	3 645	
[7, 10[	8,5	**430**	31,1%	83,9%	3 655	7 300	
[10, 13[	11,5	**102**	7,4%	91,3%	1 173	8 473	
[13, 17[	15	**70**	5,1%	96,4%	1 050	9 523	
[17, 20[		**30**					
[20, 100[		**20**					
	TOTAL:	**1382**		TOTAL:			

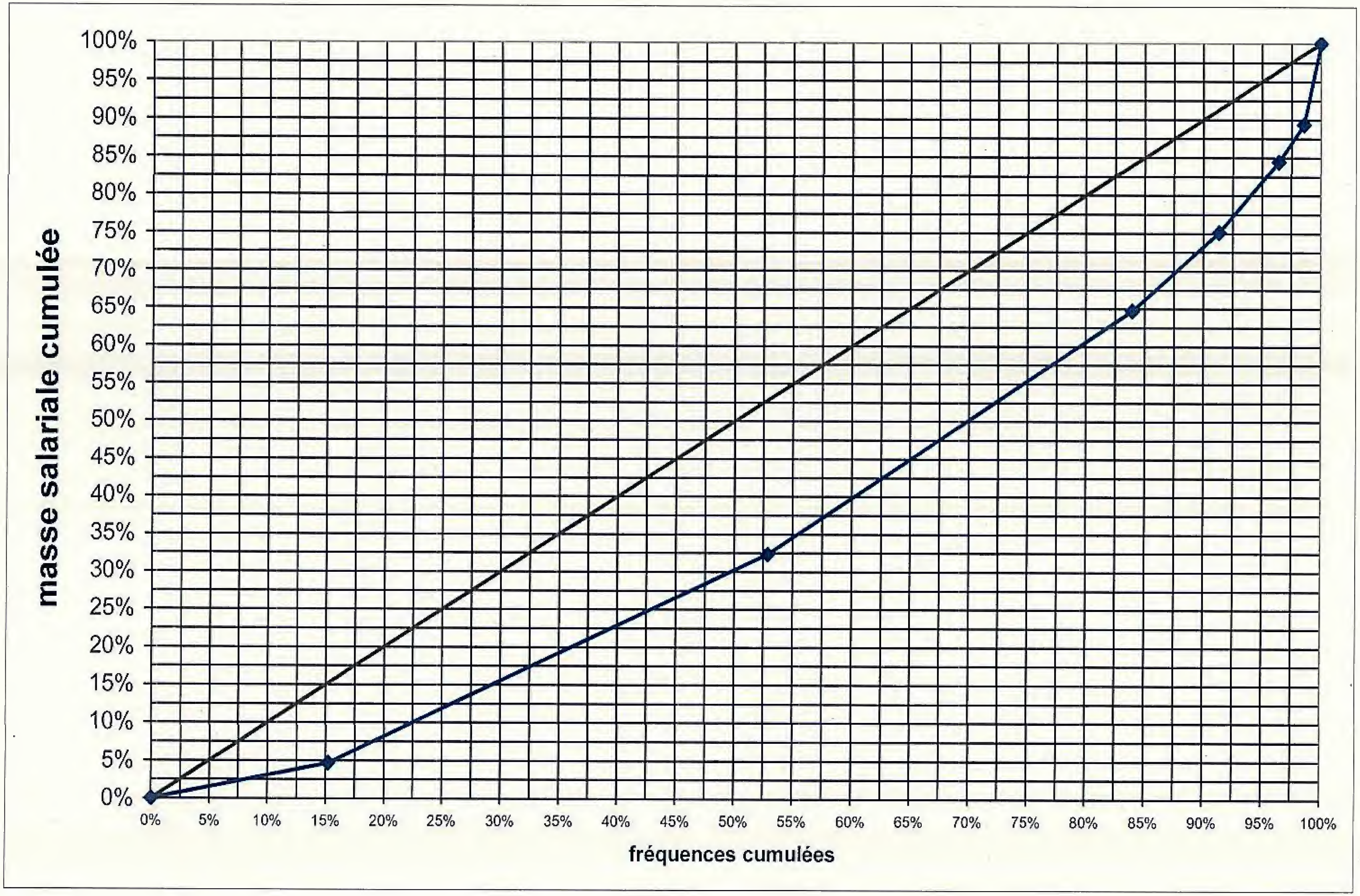
Exercice 3

Graphique des fréquences cumulées croissantes



Exercice 3

Courbe de concentration



Exercice 1:

1)

Année	Nombre d'enfants	fréquence	effectif cumulé croissant	fréquence cumulée croissante	effectif cumulé décroissant	fréquence cumulée décroissante
1	91	18,2%	91	18,2%	500	100,0%
2	72	14,4%	163	32,6%	409	81,8%
3	60	12,0%	223	44,6%	337	67,4%
4	52	10,4%	275	55,0%	277	55,4%
5	45	9,0%	320	64,0%	225	45,0%
6	40	8,0%	360	72,0%	180	36,0%
7	37	7,4%	397	79,4%	140	28,0%
8	32	6,4%	429	85,8%	103	20,6%
9	26	5,2%	455	91,0%	71	14,2%
10	25	5,0%	480	96,0%	45	9,0%
11	20	4,0%	500	100,0%	20	4,0%
TOTAL:	500					

2) 32 enfants (6,4% de la population) sont nés dans la 8e année de mariage ; 429 enfants (85,8% de la population) sont nés dans les 8 premières années de mariage et 103 enfants (20,6% de la population) sont nés à partir de la 8e année de mariage.

3) $\overline{x} = \dfrac{91\times1 + 72\times2 + \dots + 20\times11}{500} = 4,6$ ans

$V(x) = \dfrac{91\times1^2 + 72\times2^2 + \dots + 20\times11^2}{500} - 4,6^2 = 9,12$

$\sigma(x) = \sqrt{V(x)} = 3,02 \simeq 3$ ans

<u>Exercice 2</u> 1) Population = exploitations agricoles ; Individu = une exploitation ; Caractère X = âge du chef d'exploitation. Modalités = classes d'âge.

2)

Classes	Effectifs	Fréquences	Fréquences cumulées croissantes	Fréquences cumulées décroissantes
[20, 25[	580	1,2%	1,2%	100,0%
[25, 30[	2 162	4,6%	5,8%	98,8%
[30, 40[	8 063	17,2%	23,0%	94,2%
[40, 50[	9 569	20,4%	43,4%	77,0%
[50, 60[	10 660	22,7%	66,1%	56,6%
[60, 70[	15 913	33,9%	100,0%	33,9%
TOTAL :	46947	100%		

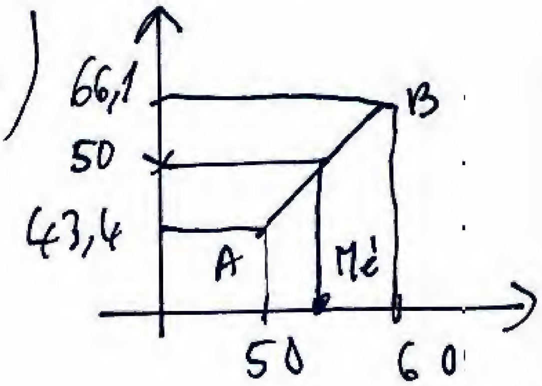
3) 77% ont au moins 40 ans ; 5,8% ont moins de 30 ans.

Moins de 60 ans → 66,1%

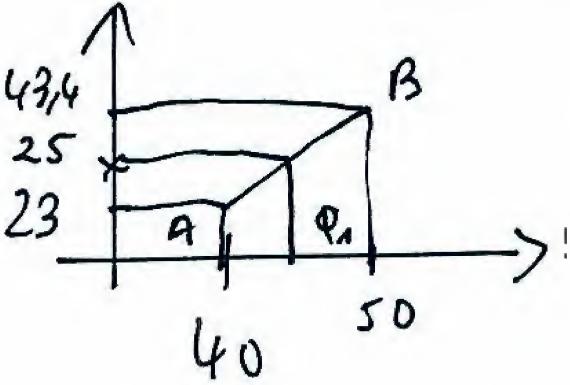
Moins de 25 ans → 1,2%

→ Entre 25 et 60 ans : 66,1% − 1,2% = 64,9%

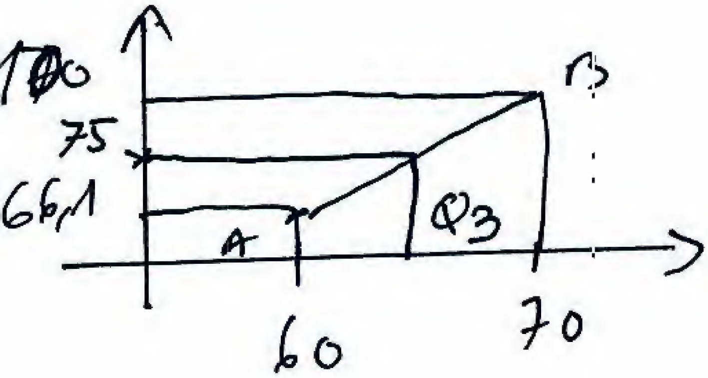
4)



Mé = 52,9 ≃ 53 ans

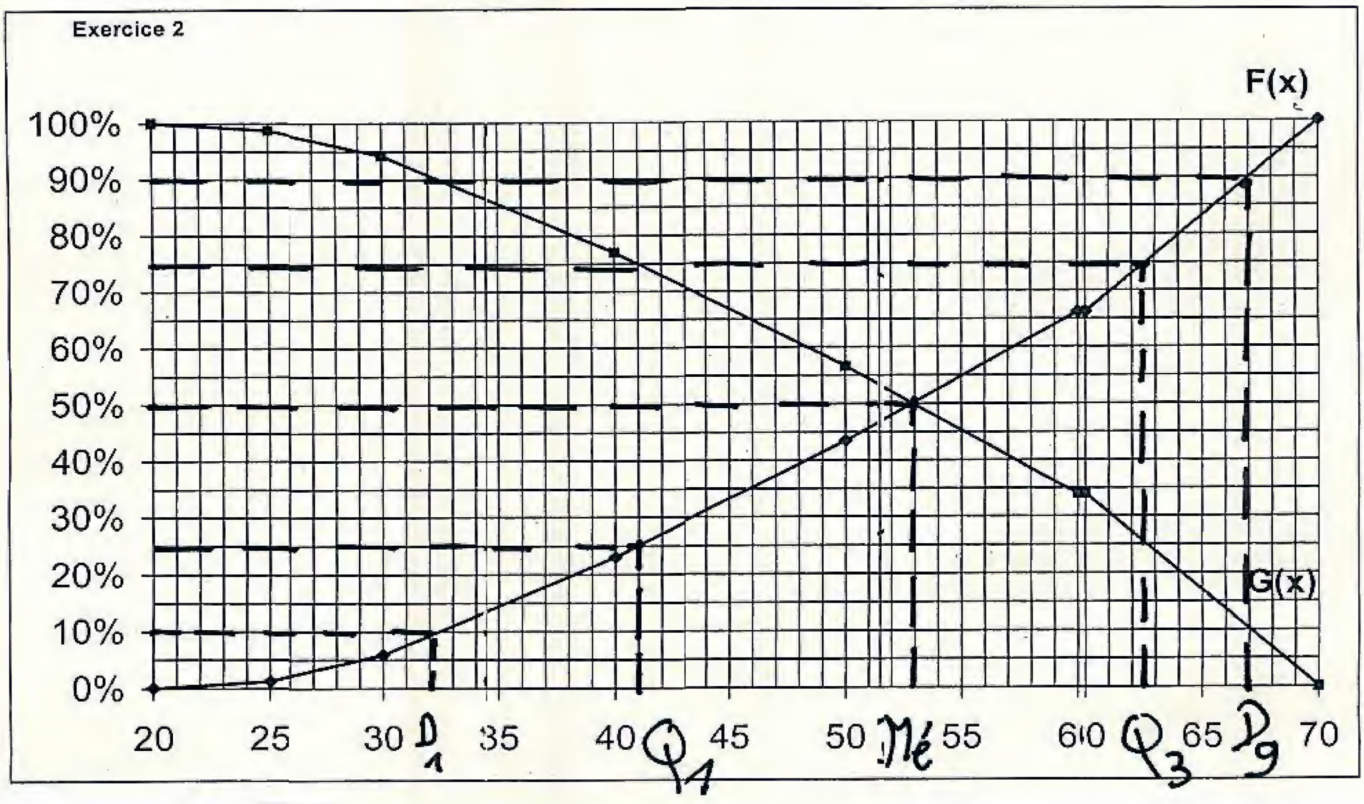


$$Q_1 = 40 + \frac{10}{20,4}(25-23)$$
$$= 41 \text{ ans}$$



$$Q_3 = 60 + \frac{10}{33,9}(75-66,1)$$
$$= 62,5 \text{ ans}$$

Graphique → $D_1 = 32$ ans $D_9 = 67$ ans



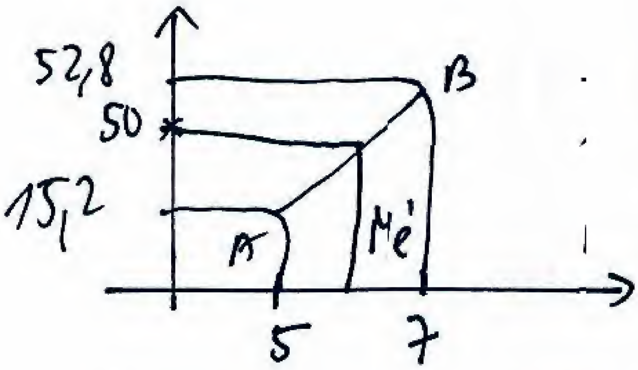
Graphique → Moins de 65 ans : 83% de la population
Moins de 35 ans : 15% de la population
→ Entre 35 et 65 ans : 68% de la population

3

Exercice 3 : 1/

Classe de salaire	centre	effectif	fréquence	fréquence cumulée	masse salariale de la classe	masse salariale cumulée	masse salariale cumulée (en %)
[0 , 5[	2,5	210	15,2%	15,2%	525	525	4,7%
[5, 7[	6	520	37,6%	52,8%	3 120	3 645	32,3%
[7, 10[	8,5	430	31,1%	83,9%	3 655	7 300	64,7%
[10, 13[	11,5	102	7,4%	91,3%	1 173	8 473	75,1%
[13, 17[	15	70	5,1%	96,4%	1 050	9 523	84,4%
[17, 20[	18,5	30	2,2%	98,6%	555	10 078	89,4%
[20, 100[	60	20	1,4%	100,0%	1 200	11 278	100,0%
	TOTAL:	1382	100,00%	TOTAL:	11 278		

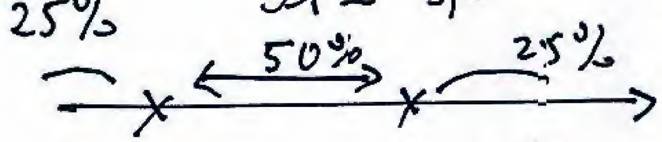
2) $\bar{x} = \dfrac{11278}{1382} = 8{,}16$ $(816 €)$



$Mé = 5 + \dfrac{2}{37{,}6}(50 - 15{,}2) = 6{,}85$

Graphique $\rightarrow$ $Q_1 = 5{,}5$ $\quad Q_3 = 9$

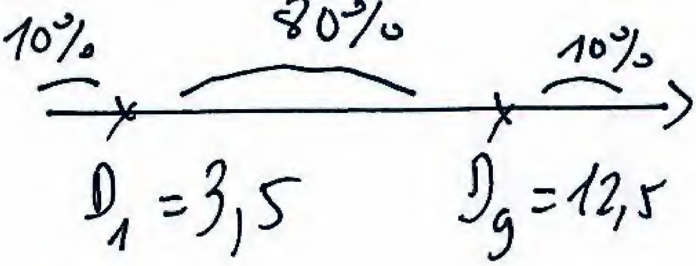
$D_1 = 3{,}5$ $\quad D_9 = 12{,}5$



$Q_1 = 5{,}5$ $\qquad Q_3 = 9$

$Q_3 - Q_1 = 3{,}5$; 50% de la population est répartie sur un intervalle de longueur 3.5

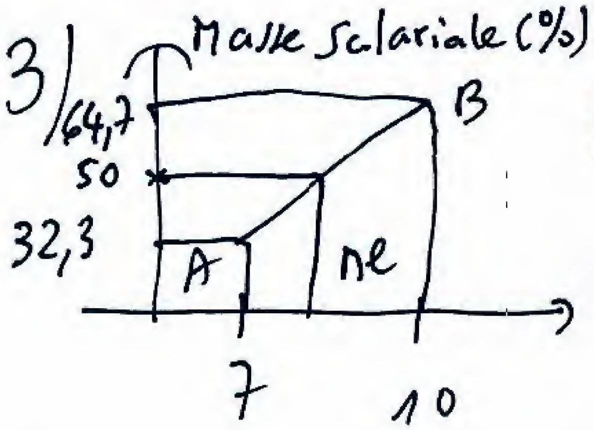
(4)



$D_9 - D_1 = 9$

80% de la population est répartie sur un intervalle de longueur 9.

3)

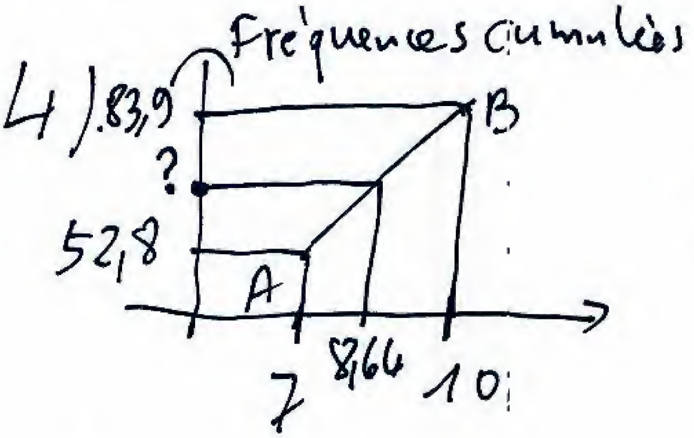


$Me = 7 + \frac{3}{32,4}(50 - 32,3)$

$= 8,64$

Ceux qui gagnent moins de 8,64 se partagent 50% de la masse salariale.

4)



$y = 52,8 + \frac{83,9 - 52,8}{10 - 7}(8,64 - 7)$

$= 69,8\% \simeq 70\%$

70% des employés se partagent 50% de la masse salariale totale.

5/ Courbe de concentration : Ceux qui gagnent moins que la médiane (soit 50% des employés) se partagent 30% de la masse salariale

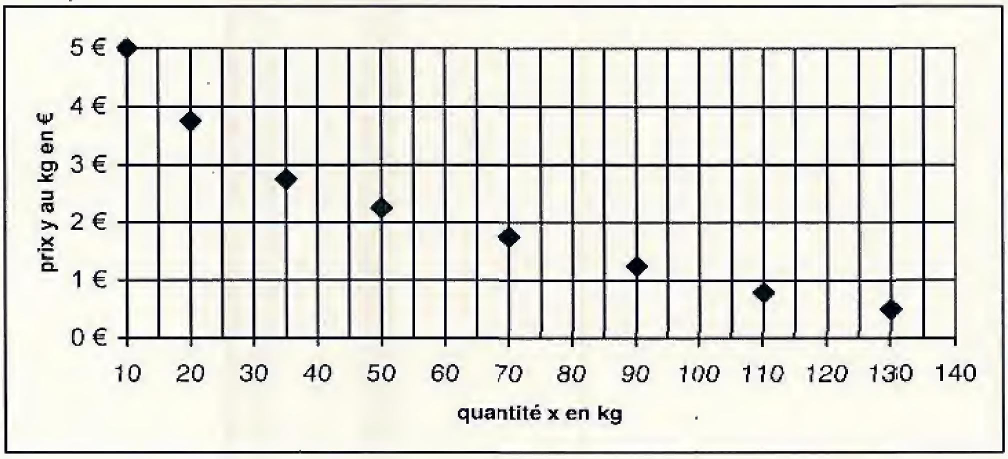
UNIVERSITE NANCY I
IUT HENRI POINCARE DE LONGWY,
DEPARTEMENT GEA

Exercices sur les ajustements linéaires et non linéaires.

Exercice 1 : L'observation des prix et des quantités sur un marché de la tomate a donné les résultats suivants :

Quantités x en kg	10	20	35	50	70	90	110	130
Prix y au kg en €	5	3,75	2,75	2,25	1,75	1,25	0,8	0,5

Ainsi, une quantité de 35 kg de tomates est vendue au prix de 2,75 € le kg.



Dans la suite de l'exercice tous les résultats obtenus à la calculatrice seront donnés avec quatre décimales.

1) Déterminer la droite d'ajustement linéaire $y = ax + b$, qui permet d'expliquer le prix au kg par la quantité achetée. Calculer le coefficient de corrélation entre x et y et expliquer son signe. Calculer le coefficient de détermination et rappeler son interprétation. Prévoir le prix d'un kg de tomates pour un achat de 140 kg. Commenter le résultat.

2) Chercher maintenant un ajustement par une fonction logarithme de la forme : $y = a\ln(x) + b$. (Indication : En posant : $u = ln(x)$ on se ramènera à un ajustement linéaire : $y = a.u + b$). Calculer le coefficient de corrélation entre u et y et le coefficient de détermination. Prévoir le prix au kg pour un achat de 140 kg.
3) Indiquer lequel de ces deux ajustements vous semble le plus judicieux (on justifiera la réponse).

Exercice 2 : Une entreprise envisage la fabrication d'un nouveau produit. Elle étudie la demande pour ce produit, afin de déterminer le prix de vente qui lui permettra de maximiser la recette.

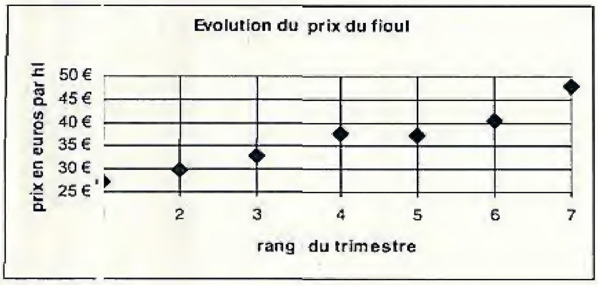
Dans le tableau suivant, figurent les résultats d'une enquête, réalisée pour déterminer la demande *y* de ce nouveau produit en fonction de son prix de vente *x* en euros.

x	200	250	300	350	450	500
y	550	430	400	310	260	210

1) Représenter graphiquement le nuage de points. Déterminer l'équation de la droite de Mayer. Placer cette droite sur le graphique
2) . Déterminer l'ajustement linéaire par la méthode des moindres carrés. Calculer le coefficient de corrélation et le coefficient de détermination.
3) On cherche maintenant à déterminer un ajustement de *y* en fonction de *x*, de la forme $y = b.x^a$, Déterminer *a* et *b* (on se ramènera à un ajustement linéaire, en posant *v = ln(y), B = ln(b) et u = ln(x).).* Calculer le coefficient de corrélation entre *u et v*, puis le coefficient de détermination. Interpréter ce dernier coefficient.
4) Estimer la demande, si le prix de vente est fixé à 400 €.
5) Lequel des deux ajustements semblent le plus judicieux ?

Exercice 3: Le tableau et le graphique ci-dessous présente l'évolution du prix du fioul domestique (en euros par hectolitre), de 2004 à 2005 .

	2004				2005		
Trimestre	1	2	3	4	1	2	3
Rang *x* du trimestre	1	2	3	4	5	6	7
Prix *y* (en euros par hl)	27,2	29,9	33	37,6	37,3	40,6	47,7



Déterminer un ajustement de y en fonction de x par une fonction exponentielle $y = b.a^x$.
En déduire une estimation du taux trimestriel moyen de hausse

Exercices sur les ajustements

Exercice 1 : 1/ $y = ax + b$. On calcule a et b par la méthode des moindres carrés. La calculatrice donne : $a = -0{,}0337$; $b = 4{,}4242$

$r = -0{,}9485 < 0$ car x et y varient en sens contraire (quand x augmente, y diminue).

Coefficient de détermination $= r^2 = 0{,}9 = 90\%$

(90% des variations de prix sont expliquées par l'ajustement linéaire).

Prévision : pour $x = 140$, $y = -0{,}0337 \times 140 + 4{,}4242$
$= -0{,}29$ €

La prévision de prix est négative, ce qui montre que l'ajustement n'est pas excellent.

2) $y = a \times \ln(x) + b$: On pose $u = \ln(x)$

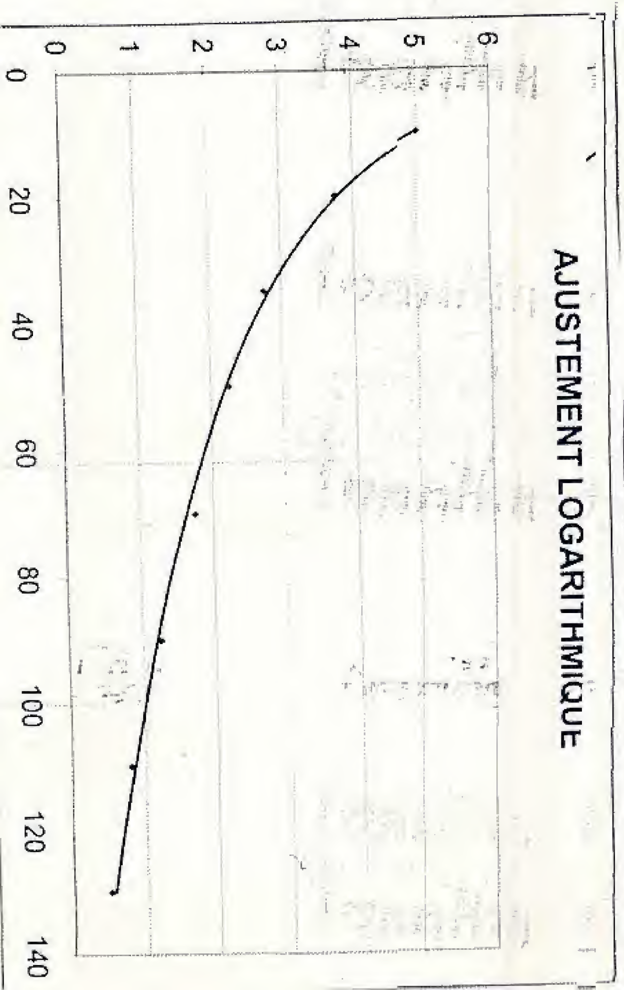
On entre dans la calculatrice les couples $u = \ln(x)$ et y.

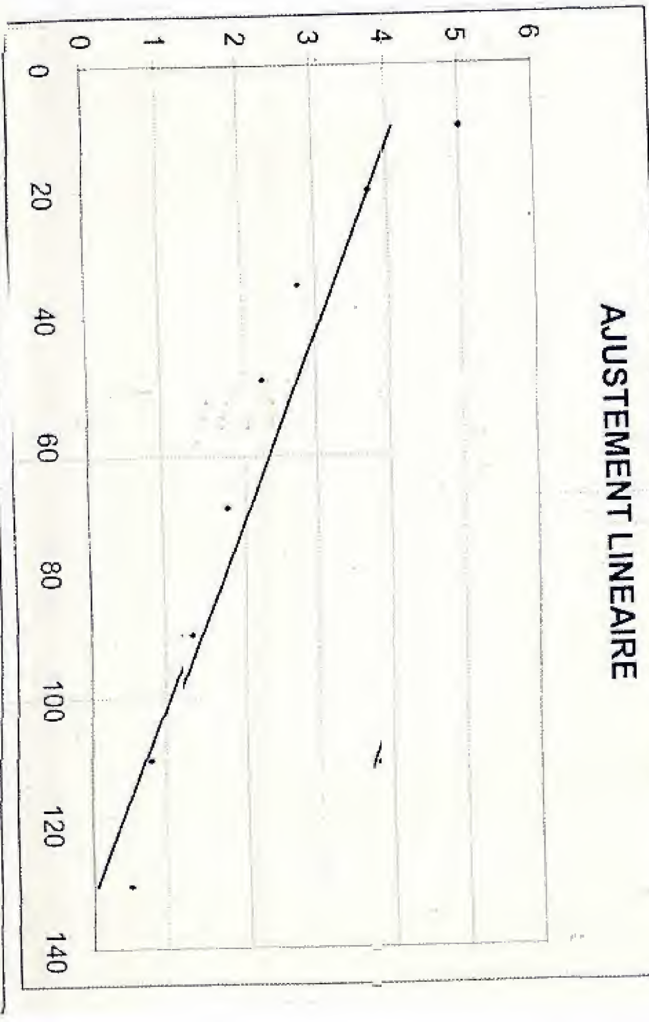
On obtient $a = -1{,}7221$; $b = 8{,}9472$

$r = -0{,}999$; $r^2 = 99{,}8\%$ → 99,8 % des variation de y sont expliquées par l'ajustement logarithmique ce qui mieux que par l'ajustement linéaire.

$$x = 140 \rightarrow y = -1{,}7221 \times \ln(140) + 8{,}9472$$
$$= 0{,}44 €$$

Pour une quantité de 140 kg, le prix prévu est de 0,44 € le kg.





Les graphiques montrent que l'ajustement logarithmique est meilleur que l'ajustement linéaire. ②

Exercice 2 : 1) On sépare le nuage en deux groupes de 3 points. On cherche le point moyen (centre de gravité) de chaque groupe.

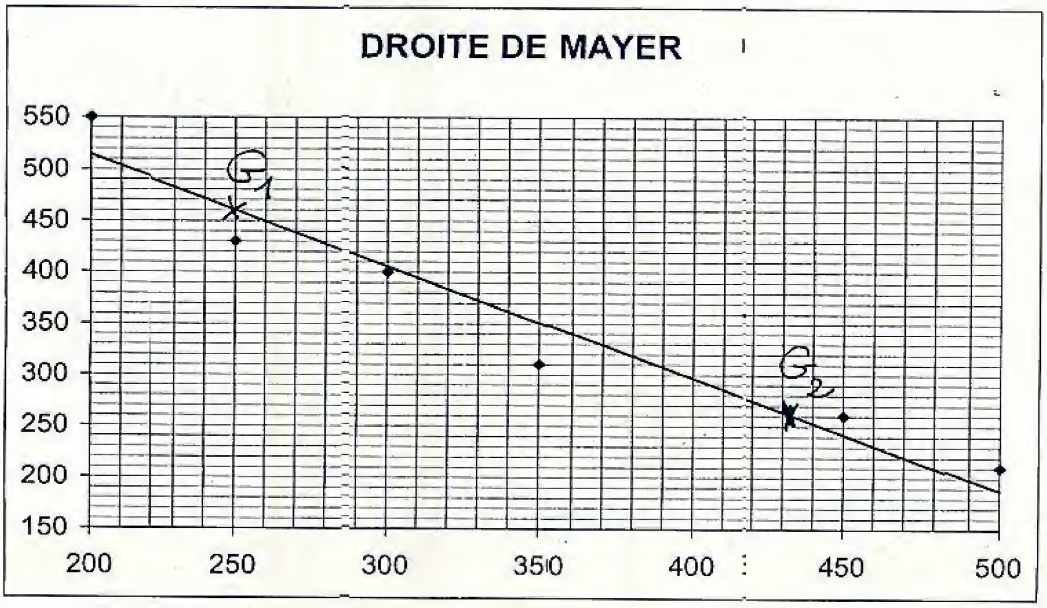
Groupe 1 : (200, 550) (250, 430) (300, 400)

$$\to G_1 \begin{cases} \frac{200+250+300}{3} = 250 \\ \frac{550+430+400}{3} = 460 \end{cases}$$

De même : $G_2 (433,33 ; 260)$.

Equation de G_1G_2 : $y = 460 + \frac{-200}{183,33}(x - 250)$

$\to y = -1,091x + 732,73$



3

Ajustement linéaire par la méthode des moindres carrés : la calculatrice donne : $a = -1{,}0435$; $b = 716{,}5217$

$r = -0{,}97 < 0$ car x et y varient en sens contraire

$r^2 = 94\%$ (94% des variation de y sont expliquées par la droite $y = ax + b$).

3) $y = b \times x^a \rightarrow \ln(y) = \ln(b) + a \times \ln(x)$

On pose $u = \ln(x)$, $v = \ln(y)$ et $B = \ln(b)$ pour obtenir : $v = a \times u + B$ (ajustement linéaire de v en fonction de u)

A la calculatrice, on entre les couples $u = \ln(x)$, $v = \ln(y)$ pour obtenir a et B.

$a = -1{,}0032$; $B = 11{,}6376$.

$B = \ln(b) \Leftrightarrow b = e^B = 113278{,}806$

On a donc $y = 113278 \times x^{-1,0032}$

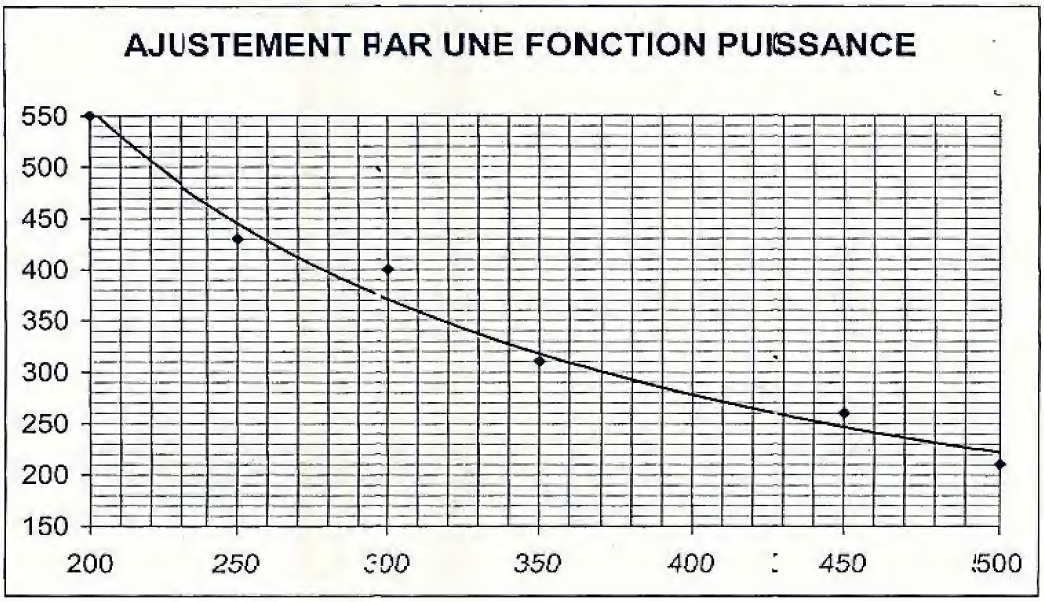
$r = -0,9891$ $r^2 = 97,8\%$. (97,8% des variation de y sont expliquées par la fonction puissance $y = b \times x^a$).

On remarque que le coefficient de détermination est plus élevé qu'avec l'ajustement linéaire.

4/ prévision pour $x = 400$

$\rightarrow y = 113278 \times 400^{-1,0032} = 278$

La fonction puissance est plus proche du nuage que la droite de Mayer.



Exercice 3 : $y = b \times a^x$. ~~[illegible]~~

$\rightarrow \ln(y) = \ln(b) + x \times \ln(a)$

On pose : $A = \ln(a)$, $B = \ln(b)$ et $v = \ln(y)$ pour obtenir $v = Ax + B$ (ajustement linéaire de v en fonction de x).

A la calculatrice, on entre les couples x et $v = \ln(y)$ pour obtenir :

$A = 0{,}0864 \quad B = 3{,}2275$

$\rightarrow a = e^A = 1{,}09$ et $b = e^B = 25{,}2$

$\rightarrow y = 25{,}2 \times 1{,}09^x$

Le coefficient multiplicateur 1,09 indique une hausse trimestrielle de 9%